...tics Statistique
...da Canada
Ottawa

PEARSONIAN TYPE I CURVE AND ITS FERTILITY PROJECTION POTENTIALS

by

S. Mitra and A. Romaniuk

...INTED PUBLICATION No. 21

...ed from Demography 10(3) August 1973, 351-365

Statistics Canada Statistique Canada

Ottawa

PEARSONIAN TYPE I CURVE AND ITS FERTILITY PROJECTION POTENTIALS

by

S. Mitra and A. Romaniuk

REPRINTED PUBLICATION No. 21

Reprinted from Demography 10(3) August 1973, 351-365

DEMOGRAPHY© Volume 10, Number 3 August 1973

PEARSONIAN TYPE I CURVE AND ITS FERTILITY PROJECTION POTENTIALS

S. Mitra
Department of Sociology and Anthropology, Emory University, Atlanta, Ga. 30322
A. Romaniuk
Statistics Canada, Tunney's Pasture, Ottawa, K1A 0T6, Canada, and University of Montreal, Montreal, Canada

Abstract—New procedures are developed in this article for estimating parameters of the Pearsonian Type I curve which are particularly adaptive to factors influencing the pattern of age-specific fertility rates. It is shown that with this model the number of parameters required for the graduation and simulation of these rates can be reduced to only three—total fertility rate, mean and modal ages of fertility. The reduction in the number of fertility parameters offers considerable operational and analytical advantages, and makes the Pearsonian Type I curve particularly appropriate for the construction of a parametric model for fertility projections. In light of the results of empirical tests based on fertility data for Canada, the model's potential for birth projections appears quite promising.

Introduction

Statisticians and demographers have shown considerable interest in the Pearsonian Type I curve, primarily as a means of graduation of age-specific fertility rates (Wicksell, 1931; Mitra, 1967; Keyfitz, 1968; Avery, 1970). Although a few authors have also recognized its potential for birth projections (Tekse, 1967; Stone, 1970), to the best of our knowledge no one has yet systematically investigated this aspect. This article has two objectives: first, to simplify the procedures for estimating the parameters of the equation of the Pearsonian Type I curve, which is employed in the graduation of age-specific fertility rates; and second, to investigate the feasibility of utilizing this curve as a foundation in the construction of a parametric model for birth projections. The two objectives are closely related; the first objective is in fact a prerequisite of the second. However, a model which is satisfactory for graduation purposes might be unsatisfactory for projections. For example, higher order moments may help to perfect the graduation, but they are often inappropriate for forecasting because they are difficult to interpret in demographic terms. One of the high-priority goals of demographers is to develop a projection model of increased operational and analytical capability. However, if this goal is to be attained, it is essential that a procedure be developed which would rely on a limited number of demographically meaningful fertility parameters, and simultaneously, one which would permit age-specific fertility rates to be satisfactorily reproduced. The derivation of a procedure which would satisfy these two criteria is the ultimate goal of this article.

In the first section, the Pearsonian Type I curve is briefly presented, and a justification is provided for its selection as a mean of reproducing age-specific fertility rates. In the second section, there is a short review of the procedures

usually used for deriving constants in the equation of the Pearsonian Type I curve, followed by a presentation of new methods which have been developed to estimate the constants. These new procedures are tested in the third section by using Canadian historical data on fertility, and in the final section, attempts are made to explore the ways in which the Type I function could be utilized for birth projections.

Presentation and Justification of the Type I Curve

With the origin at the mode, the equation of the Type I curve is:

$$y = y_o(1 + x/a_1)^{m_1}(1 - x/a_2)^{m_2}, \qquad (1)$$

where y_o is the modal ordinate and $-a \leq x \leq a_2$.
Also,

$$m_1/a_1 = m_2/a_2. \qquad (2)$$

In formula (1), y_o is the modal fertility rate, which in human populations falls between 20 and 30 years of age; a_1 and a_2 jointly determine the reproductive interval; and m_1 and m_2 determine the shape of the fertility curve. When m_1 and m_2 are approximately equal, the curve approaches a normal distribution, but it is positively skewed when $m_1 < m_2$. Modern reproductive patterns of population are characterized by the latter skewed distribution; this is demonstrated in Appendix Table A-1, which presents the parameters of a fertility distribution obtained from data for Canada. One can see, as a first indication of skewness, that the modal age is consistently smaller than the mean age, and that the measure of skewness, β_1, is positive throughout the 1926–1969 period.

However, the selection of a particular Pearsonian curve to describe the fertility experience of any given country cannot be based merely on impressionistic views and judgments based on the parameters mentioned above. Indeed, K. Pearson (Elderton, 1930) has argued that the criterion for this selection should be based on the calculation of k:

$$k = \beta_1(\beta_2 + 3)^2/[4(2\beta_2 - 3\beta_1 - 6) \cdot(4\beta_2 - 3\beta_1)].$$

Pearson recommends that the Type I curve should be used when k has a negative value.

It can be seen in Appendix Table A-1 that for Canada k has a negative value, and consequently, that the selection of this curve is justified here. It should be noted, however, that k assumes progressively higher negative values. In view of these shifts in the age pattern of Canadian fertility, it would be worthwhile trying some of the other curves in the Pearsonian system of curves, such as the "normal curve" or its proxy, Type II, to represent the age pattern of fertility in earlier years and the transitional Type III curve to represent the age pattern of fertility in more recent years. Alternatively, outside of models based on the so-called gamma distribution, only the Mazur-type model (Mazur, 1963) has proved to be appropriate in the case of relatively high and less asymmetrical fertility (Tekse, 1967).

A comparative study of fitting various curves is outside the scope of this paper. Instead, we shall concentrate our attention on the curve which, based on the criterion of k, appears to be most suitable for the period under observation and which, by implication, will probably continue to conform to Canada's fertility patterns in the near future. Tests conducted by Mitra (1967) and Avery (1970), using fertility materials for nations throughout the world, have revealed that the Pearsonian Type I function is among the most suitable mathematical functions developed so far for approximating observed age schedules of fertility. Unfortunately, in view of data constraints, their fitting of the curve had to be restricted to fertility rates by five-year age groups, and consequently a de-

tailed analysis of the goodness of fit was not possible.

Our next task is to discuss the various methods by which the constants in formula (1) can be derived.

Methods of Deriving Constants

The method most commonly used for fitting formula (1) to the actual age-specific fertility rates is the method of moments. However, the number of moments used depends on the particular procedure followed. In the Elderton (1930) procedure, the constants a_1, a_2, m_1, and m_2 in equation (1) are calculated from the first four moments of the frequency distribution, that is, from the mean, variance, skewness, and kurtosis. Later, Mitra (1967) developed a procedure that reduces the required number of moments to the first two, but this procedure assumes a fixed age interval of fertility. Thus, with the fertility interval set at ages 15 through 50,

$$a_1 + a_2 = 35. \quad (3)$$

With equations (2) and (3), the required number of independent parameters is reduced to only two, and the solutions are obtained from the following equations:

$$m_1 + m_2 = \mu_1'(a_1 + a_2 - \mu_1')/\mu_2 - 3 \quad (4)$$

and

$$m_1 = [(m_1 + m_2 + 2)/(a_1 + a_2)]\mu_1' - 1, \quad (5)$$

where μ_1' represents the mean counted with the starting point at age 15 and μ_2, the variance. Equations (2) and (3) can then be used to determine a_1 and a_2. As for y_o, the following can be applied:

$$y_o = N/[(a_1 + a_2)B(m_1 + 1, m_2 + 1)] \cdot [m_1^{m_1} \cdot m_2^{m_2}]/[(m_1 + m_2)^{m_1+m_2}]. \quad (6)$$

In equation (6), N represents the sum of the age-specific fertility rates and B represents the Beta function. In fact, y_o may be regarded as a multiplier that equalizes the sum totals of the observed and graduated distributions.

The above procedure is based on equating the first two moments of the observed distribution with those of the theoretical distribution. Alternatively, the procedure for calculating the constants may be further simplified by using the first moment of the distribution, the mean, and some other measure, such as the mode. A conventional method may be utilized to estimate the mode, or one could simply substitute the midpoint of the single-year age interval with the highest age-specific fertility rate. In this case, given the value of the mode or a_1 with the origin at the start of the curve, the remaining parameters can be obtained from the following:

$$a_2 = (a_1 + a_2) - a_1, \quad (7)$$

$$m_2 = a_2(a_1 + a_2 - 2\mu_1')/[(a_1 + a_2) \cdot (\mu_1' - a_1)], \quad (8)$$

and

$$m_1 = (a_1/a_2)m_2. \quad (9)$$

Although the reduction in the number of moments may entail some loss in goodness of fit, it presents some definite analytical and operational advantages. These advantages will be examined in greater detail later, but at the present time let us merely emphasize that the model's parameters should respond in a manner that is consistent with the trend in fertility rates. In the absence of this consistency, it would not be possible to relate the variations in the distributions to the variations in the parameters. Consequently, the model would be of little practical value. Given the multitude of possible parameters, it is extremely difficult to relate, for example, the effect of a reduction in the total fertility rate on the modal age, the fertile age span, etc. The restrictions on the fertility interval in the second procedure and on the fertility interval and modal age in the third procedure make the interpre-

tation of the remaining parameters relatively simple and perhaps more meaningful.

Other procedures for deriving or perfecting the estimation of the constants may be suggested. One such procedure, based on the fertile age range, the modal age, and the modal fertility rate, also seems reasonable and worthy of investigation. With the origin at the start of the curve, and using the relationship

$$m_1/a_1 = m_2/a_2$$
$$= (m_1 + m_2)/(a_1 + a_2) = C, \quad (10)$$

equation (6) for the modal fertility rate can be rewritten as:

$$y_o = \frac{N}{(a_1 + a_2)B(Ca_1 + 1, Ca_2 + 1)} \cdot \left[\left(\frac{a_1}{a_1 + a_2}\right)^{a_1}\left(\frac{a_2}{a_1 + a_2}\right)^{a_2}\right]^{c} \quad (11)$$

where N for the single-year distribution is equivalent to the total fertility rate. The constant C, obtained through an iterative procedure, can then be used in (10) to estimate the remaining parameters.

A different approach that could be used to improve the fit has been suggested by Keyfitz (1968), namely, iteration techniques based on a descent method. This approach proved useful in fitting the Gompertz curve to Canadian data on cumulative fertility (Murphy and Nagnur, 1972), and with due adaptation it could be applied to fertility data in the form of age-specific distributions.

The two latter procedures are mentioned only to indicate further developments to improve the fit. At this point we were not able to fully investigate their potential for projections. Instead we have concentrated on the first three procedures and have tested them against empirical data, the results of which are to be presented in the next section.

For simplicity in future references, these three procedures will henceforth be referred to by the following codes shown in parentheses:

a) Method based on the first four moments (4M)
b) Method based on the first two moments (2M)
c) Method based on the mean and the mode (1M).

The numerals 4, 2, and 1 stand for the number of moments on which a particular method is based.

Empirical Verifications of the Procedures and the Model

Whatever may be the theoretical justification of a particular method, any decision concerning its suitability for practical application must rest on some empirical test, such as the comparison of derived estimates with actual figures. Before proceeding with the testing, a brief discussion of the data base is in order.

These tests are performed using data on fertility by single years of age for Canada available since 1926. Except for the personal affiliation of one of the authors, there is no particular reason for selecting this country. Canadian data on fertility are assumed to be fairly good, but not completely free from reporting errors. This is probably true, particularly for the earlier years which, among other things, are characterized by gross irregularities in the modal age group. There seems to be no theoretical argument in favor of the oscillations observed at alternate ages in this age group. The extent to which the remaining ages have been affected by this kind of error is not known, and no attempt has been made to adjust the data for such biases. However, the magnitude and direction of these biases are not expected to greatly influence the characteristics of the distributions on which the estimates of the parameters are based.

Only period fertility data are used here. Data by cohorts are reconstructions

from the period data, supplemented by extrapolations from truncated cohorts. While these cohort data are useful for general types of analysis, as such they do not warrant enough confidence to base on them the assessment of the validity of the procedures being tested here.

For convenience of space, the years 1926, 1931, 1941, 1951, 1961, and 1969 have been selected for statistical comparison. During this time period there were considerable variations in both the level and age pattern of fertility. In this paper, estimates for each year within the 1926–1969 period are provided only for method (4M) (Appendix, Table A-2). A priori, one may expect the closeness of fit to improve as the number of moments used is increased. It follows that method (4M) sets some standard against which estimates obtained by other methods can be gauged.

We shall compare first the estimated fertility measures and parameters with the actual ones, and then examine the goodness of fit for the overall fertility age schedule, obtained by the three procedures described in the preceding sections.

Estimated and Actual Fertility Measures

Of all the parameters of the Type I distribution, the one that is most meaningful in this context seems to be the modal age. Since method (2M) is restricted by a fixed fertility interval, the estimate of the modal age will not be unique and will depend on the specific choice of that interval. The results based on two alternative fertility intervals of 15–50 and 17–50 years by (2M) along with those calculated by (4M) are compared with observed modal ages in Table 1. It should be noted that the observed values presented in the table were ad-

TABLE 1.—Estimates of Modal Age for Canada Using Methods (4M), (2M), and (1M): 1926, 1931, 1941, 1951, 1961, and 1969

Year	Modal Age Obtained from Method				Total Fertility (per 1000)
	(4M)	(2M) (15-50)[a]	(2M) (17-50)[a]	(1M) (observed)	
(1)	(2)	(3)	(4)	(5)	(6)
1926	28.1 (17.0-48.5)[b]	28.9	27.9	28.0	3356
1931	27.6 (16.9-49.3)[b]	28.5	27.6	27.7	3201
1941	26.3 (17.2-49.6)[b]	27.6	26.4	27.1	2824
1951	25.5 (17.0-49.9)[b]	26.6	25.4	25.3	3480
1961	24.2 (17.2-50.4)[b]	25.7	24.3	24.0	3857
1969	24.0 (16.6-53.0)[b]	25.1	23.7	23.8	2410

a-Fixed fertility interval.
b-Estimate of the fertile age range based on (4M).

justed for irregularities which appear to have been caused by random fluctuations and age misreporting. These adjustments were made by tracing a freehand curve through the points of observed modal ages. It should also be recalled that method (1M) has no built-in mechanism for the calculation of modal age, and that the latter has to be given in order to enable us to use this method for deriving other parameters. Methods (4M) and (2M), with an age interval of 17–50, estimate the modal age fairly well, whereas method (2M), with a fertility range of 15–50, tends to overestimate it by about one year.

The figures in parentheses in column (2) are the estimates of the fertile age range based on (4M), and it can be seen that the lowest (16.6) and the highest (53.0) limits occur in 1969. The lower limit is no larger than 17.2 years, and the smallest value of the upper limit is 48.5 years. The latter appears to vary more than the former. In general, the derived measure of reproductive period is not quite in accord with the actual one observed among human populations. Such inconsistencies are unavoidable when all the parameters are estimated from the moments of distribution.

As for the modal fertility rate shown in Table 2, no clear pattern emerges from the differences between the estimates obtained by various methods. They all appear to estimate modal fertility fairly well, with the exception of (1M) (15–20), which tends to understate it considerably.

Parameters m_1 and m_2, which were derived by various procedures, are shown in Tables 3 and 4. As could be expected, the values of these parameters are, and should be, independent of the total frequency since their derivation is dependent primarily upon the distribution of the relative frequencies. Accordingly, changes in the parametric values are reflections of the changes in the pattern of the distributions themselves. Parameter m_2 appears to be particularly sensitive to such changes. This is reflected in the high coefficients of correlation shown in Table 5, which indicates the association between m_2 and various parameters of the age pattern of fertility, such as mean and modal ages, variance, skewness, and kurtosis. Over the period of observation, m_2 has exhibited a rather steady upward trend, which merely seems to reflect the downward shift in the age pattern of fertility in Canada. During the same period, when one disregards relatively

TABLE 2.—Estimates of Modal Fertility Rate for Canada Using Methods (4M), (2M), and (1M): 1926, 1931, 1941, 1951, 1961 and 1969

	Modal Fertility Rate Obtained from Method					Observed Modal Fertility Rate
		(2M)		(1M)		
Year	(4M)	(15-50)[a]	(17-50)[a]	(15-50)[a]	(17-50)[a]	
1926	178	181	181	152	184	188
1931	174	175	175	150	181	183
1941	159	159	160	147	173	168
1951	205	203	205	172	204	210
1961	239	231	236	195	231	257
1969	157	150	155	130	155	168

a-Fixed Fertility interval.

TABLE 3.—Estimates of m_1 for Canada Using Methods (4M), (2M), and (1M): 1926, 1931, 1941, 1951, 1961, and 1969

	m_1 Obtained from Method				
		(2M)		(1M)	
Year	(4M)	(15-50)[a]	(17-50)[a]	(15-50)[a]	(17-50)[a]
1926	.90	1.55	.98	.83	1.05
1931	.95	1.55	.97	.89	1.09
1941	.83	1.49	.89	1.13	1.23
1951	.85	1.45	.82	.75	.81
1961	.70	1.37	.73	.64	.64
1969	.95	1.39	.72	.76	.74

a-Fixed fertility interval.

minor fluctuations (see Appendix, Table A-2), m_1 has remained relatively stable; apparently, it has been less influenced by the changing age pattern of fertility. However, its apparent high (negative) correlation with the total fertility rate, shown in Table 5, is difficult to explain.

In light of the results presented in Tables 3 and 4, it is also quite apparent that m_1 and m_2 must be very sensitive to the estimation procedure. The estimates obtained by the methods differ among themselves and in relative terms; the magnitude of these differences are considerable.

Goodness of Fit

So far, attention has been focused on the agreement between the observed and estimated fertility parameters. We shall now see to what extent the derived fertility age distribution agrees with that of the observed.

The classical approach of measuring goodness of fit is to calculate X^2, the magnitude of which depends in part upon the relative difference between the observed and graduated values. This method was considered inappropriate for a situation characterized by a large number of intervals, for which some of the

TABLE 4.—Estimates of m_2 for Canada Using Methods (4M), (2M), and (1M): 1926, 1931, 1941, 1951, 1961 and 1969

	m_2 Obtained from Method				
		(2M)		(1M)	
Year	(4M)	(15-50)[a]	(17-50)[a]	(15-50)[a]	(17-50)[a]
1926	1.67	2.35	2.00	1.40	2.11
1931	1.93	2.46	2.08	1.57	2.28
1941	2.12	2.68	2.25	2.14	2.82
1951	2.43	2.92	2.43	1.81	2.40
1961	2.64	3.12	2.57	1.86	2.39
1969	3.71	3.42	2.80	2.25	2.84

a-Fixed fertility interval.

TABLE 5.—Correlation Matrix

	a_1	a_2	m_1	m_2	Total Fertility	Modal Fertility	Modal Age	Mean Age	Std. Deviation	Skewness	Kurtosis	k	Δ
a_1	1.00	-0.94	0.84	-0.81	-0.49	-0.75	0.95	0.97	-0.93	-0.95	-0.92	-0.93	-0.67
a_2	-0.94	1.00	-0.63	0.96	0.24	0.57	-0.96	-0.99	-0.98	0.98	0.99	0.99	0.67
m_1	0.84	-0.63	1.00	-0.38	-0.72	-0.83	0.69	0.70	0.60	-0.68	-0.58	-0.63	-0.61
m_2	-0.81	0.96	-0.38	1.00	0.02	0.36	-0.88	-0.91	-0.95	0.92	0.97	0.95	0.59
Total Fertility	-0.49	0.24	-0.72	0.02	1.00	0.93	-0.42	-0.36	-0.26	0.31	0.21	0.24	0.26
Modal Fertility	-0.75	0.57	-0.83	0.36	0.93	1.00	-0.71	-0.66	-0.58	0.62	0.55	0.57	0.47
Modal age	0.95	-0.96	0.69	-0.88	-0.42	-0.71	1.00	0.98	0.96	-0.96	-0.96	-0.96	-0.63
Mean age	0.97	-0.99	0.70	-0.91	-0.36	-0.66	0.98	1.00	0.98	-0.98	-0.98	-0.98	-0.65
Std. deviation	0.93	-0.98	0.60	-0.95	-0.26	-0.58	0.96	0.98	1.00	-0.98	-0.99	-0.98	-0.61
Skewness	-0.95	0.98	-0.68	0.92	0.31	0.62	-0.96	-0.98	-0.98	1.00	0.98	0.99	0.70
Kurtosis	-0.92	0.99	-0.58	0.97	0.21	0.55	-0.96	-0.98	-0.99	0.98	1.00	0.99	0.66
k	-0.93	0.99	-0.63	0.95	0.24	0.57	-0.96	-0.98	-0.98	0.99	0.99	1.00	0.70
Δ	-0.67	0.67	-0.61	0.59	0.26	0.47	-0.63	-0.65	-0.61	0.70	0.66	0.70	1.00

(4M) estimates were zero, namely, those at the beginning and the end of the fertility curve. By comparing the observed and expected distributions, one can obtain overall index values in a number of ways. One such index, the index of dissimilarity ($\triangle$), is obtained by reducing the two distributions to percentage form and summing only the positive differences between corresponding percentages. A given value of this index indicates the percentage of observations which must be redistributed among intervals in order that the two distributions become identical. It is easy to see that $\triangle$ can range from zero to 100 and that its magnitude depends on the choice of class intervals.

In Table 6, the overall error appears to be relatively small. As could be expected, the index values are consistently smallest for the (4M) method, but considering the simplicity and logical consistency of the other methods, the values are not very high. Hence, in general the model seems to be quite satisfactory, at least as a first approximation of a graduation formula for fertility rates. Only method (1M), with a fertility age interval of 15 to 50, results in more consequential error.

In order to obtain some indication of the kind of age-associated biases generated by the model, the deviations of the derived from the observed age-specific fertility rates have been calculated and depicted in Figure 1. The outstanding feature of this figure is the fact that for all observed years there is a very definite pattern in the age biases. The model tends to alternate between underestimating and overestimating the fertility rates for a series of successive age categories. Similar cyclical patterns of age biases are generated by the other two procedures.

If the model is to be used for projections, it is important to ascertain the structural conditions of fertility for which the model performs best. To this effect, correlation coefficients between the index of dissimilarity and the measures of the age pattern of fertility were calculated. As Table 5 reveals, all of the correlation coefficients are significant, but the highest coefficient (.70) is the one obtained with criterion k, which can be viewed as an index of the combined effect of skewness and kurtosis on the fertility curve's overall shape. It follows that as k increases, that is, as its negative values become increasingly larger, the error generated by the model becomes progressively larger. Consequently, after k reaches a certain level, an alternative curve such as the transitional Type III curve mentioned earlier may

TABLE 6.—$\triangle$ Values for Canada Using Methods (4M), (2M), and (1M): 1926, 1931, 1941, 1951, 1961 and 1969

	$\triangle$ Values Obtained from Method				
		(2M)		(1M)	
Year	(4M)	(15-50)[a]	(17-50)[a]	(15-50)[a]	(17-50)[a]
1926	2.74	4.08	2.77	8.94	2.80
1931	2.66	3.78	2.59	7.15	2.27
1941	2.32	4.30	2.39	5.33	4.51
1951	2.28	4.03	2.41	9.42	2.41
1961	2.97	5.38	3.30	11.34	4.03
1969	3.48	5.23	3.99	10.39	4.13

a-Fixed fertility interval

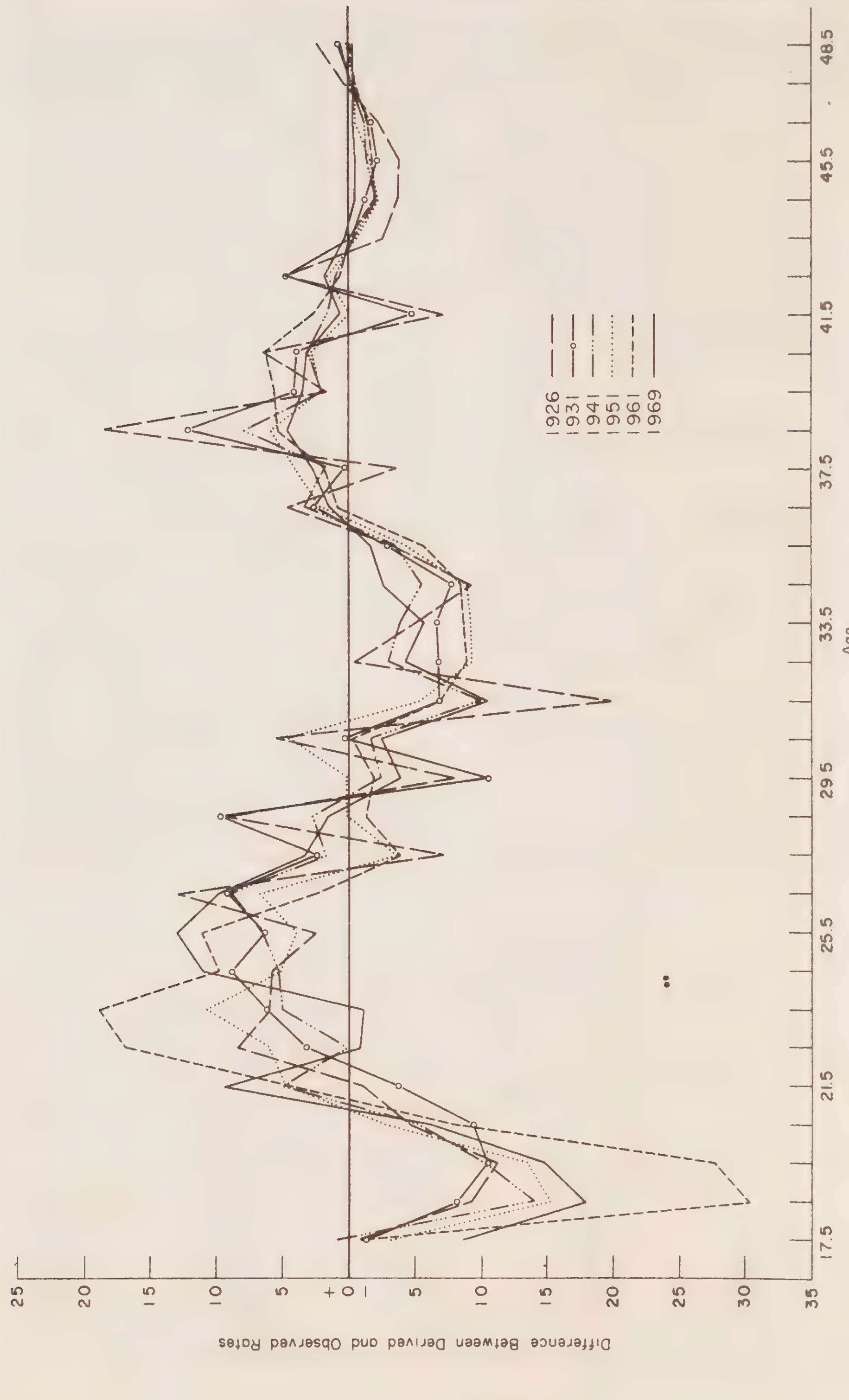

Fig 1 — Absolute Deviation of Derived (4M) from Observed Age Specific Rates for a Few Selected Years for Canada

provide a better fit than the Type I curve. However, the negative values for k which were obtained from Canadian fertility data are not nearly large enough to justify the utilization of the Type III curve. The errors in the results which are generated by each of the three methods differ from each other in magnitude, but not in pattern. Therefore, although the coefficients discussed here refer to the results derived by method (4M), the inferences to which they give rise apply to each of the methods.

Potentials for Birth Projections

In the preceding sections, the Type I curve has been examined as a device for the graduation or reproduction of observed age-specific fertility rates. In the present section, an attempt will be made to extend its application to the field of projections. We shall first demonstrate how this can be done, and then illustrate how the operational and analytical advantages of the model based on this curve make it a more effective device for projecting birth series than the conventional methods.

In trying to adapt the Type I curve to projection requirements, it is helpful to distinguish between parameters which refer to the *level* of fertility and those which refer to the *age distribution* of fertility. It should be recalled that in formula (1) the former are represented by the modal fertility rate, y_o, and the latter by the parameters a_1, a_2, m_1, and m_2.

The fertility *level* may be projected in one of two ways. One could project the modal fertility rate, in which the model in its formulation by Pearson is stated, or alternatively, one could project the total fertility rate. Although there is a very high correlation (.93) between the modal and total fertility rates, on statistical and analytical grounds it seems preferable to project the latter rates. This is because the historical series of modal fertility are influenced by irregularities which are probably attributable to age misreporting and random fluctuations. Moreover, the utilization of total fertility rates permits one to conduct analysis in much more depth than is possible with the modal fertility rate. For example, the cohort approach can be employed to explain variations in the total fertility rate by referring to changes in family size and shifts in the birth timing of successive cohorts of women (Ryder, 1969). Furthermore, future fertility trends can be forecast by analyzing parity distribution data and information obtained from surveys on intended family size (Siegel and Akers, 1969). These advantages of employing the total fertility rate as the basis for projecting future fertility levels cannot be attained if one attempts to make projections in terms of modal fertility; in this latter case the forecaster can rely only on past trends as a guide to the future.

Let us now turn our attention to the projection of fertility age distributions. In this case one can directly project either the *dependent* parameters, a_1, a_2, m_1, and m_2, or the *independent* parameters, that is, the moments and other measures which are associated with fertility frequency distributions. The selection of one of the three retained procedures as the basis for calculating the dependent parameters determines the choice of the appropriate independent parameter.

In our view, there are primarily two prerequisites which must be met if a particular set of parameters is to be selected for incorporation into a model for projections, and these are analytical suitability and amenability to a demographically meaningful interpretation. Thus, a model which yields a "good fit" but does not simultaneously meet the prerequisites mentioned here may be satisfactory for graduation purposes, but it would be totally inadequate for projection purposes. According to these criteria, *dependent* parameters which provide mathematical descriptions but do not provide meaningful demographic

interpretations do not qualify for selection. At this point the difficulties we have encountered earlier in this article in offering a demographically meaningful interpretation of the behavior of m_1 and m_2 should be recalled. To a certain extent, moments of higher orders suffer a similar fate.

Among the three listed procedures, the one which, in our opinion, best satisfies the criteria of "rationality" and demographic meaningfulness of parameters is method (1M), which requires the knowledge of merely two simple and relatively easily understood fertility measures, namely, mean and modal ages of fertility. Moreover, since these two measures are highly correlated (.98), for all practical purposes one may confine the analysis and projection to only one of them. Whereas it is true that procedure (1M) does not provide quite as close a fit as do other procedures such as (4M), the resulting loss in precision is more than compensated for by the increased operational and analytical advantages which are possible with the former method.

The accuracy of a projection based on any particular method can only be determined *ex post facto*. However, for the sake of argument, let us momentarily assume that all the *independent* parameters had been correctly projected for the 1926–1970 period, and that each of the models developed in this paper had been used to reproduce the total annual numbers of births during this period. Given these assumptions, any differences between the actual totals of annual births and those derived from the models must necessarily be due to the models. The ratios of the number of births estimated by the different models to the actual numbers of births for Canada in selected years are shown in Table 7. One can see that the ratios are all close to unity. This is true even for method (1M), which relies on only three fertility measures, namely, the total fertility rate and the mean and modal ages of fertility. Consequently, in the calculation of the total annual number of births in the indicated period, each of the three models performs almost to perfection.

Conclusion

Additional research is needed before a parametric model of fertility projections of the type outlined here can be made fully operational. This paper contains, nevertheless, important components for such a model. It is shown that fertility distribution by age can be derived by

Table 7.—Ratio of Estimated to Actual Number of Annual Births for Canada Using Methods (4M), (2M), and (1M): 1926, 1931, 1941, 1951, 1961, and 1969

	Ratio of Estimated to Actual Number of Births Obtained from Method				
		(2M)		(1M)	
Year	(4M)	(15-50)[a]	(17-50)[a]	(15-50)[a]	(17-50)[a]
1926	1.001	.986	.999	1.005	1.000
1931	1.001	1.001	1.000	1.010	.998
1941	.998	.997	.998	1.000	.996
1951	.998	.997	.997	.984	.996
1961	1.003	1.005	1.001	1.013	1.002
1969	1.008	1.006	1.011	1.018	1.009

a-Fixed fertility interval

mathematical functions from only a limited number of parameters that need to be projected. The paper offers alternative ways of calculating these parameters; it tests their results against fertility data for Canada; and it points to potentials for further developments in the areas of fertility projections. Reduction of the age schedule of fertility to only a few meaningful parameters makes possible an in-depth analysis to a degree that cannot be achieved with conventional procedures making direct use of age-specific fertility rates in projecting births. Furthermore, the model is a powerful labor-saving device, for in defining fertility in terms of a mathematical function, computers can be utilized to perform many of the involved operations.

APPENDIX

TABLE A-1.—Parameters of Canadian Fertility Distribution by Age, 1926 to 1969

Year	Total Fertility Rate (Per 1000)	Observed Modal Fertility	Modal Age[a]	Mean Age of Fertility	Std. Deviation	Measure of Skewness β_1	Measure of Kurtosis β_2	Criterion κ
1926	3360	188	27.98	30.13	6.61	.059	2.285	-0.029
1927	3320	181	27.94	30.11	6.62	.059	2.284	-0.029
1928	3300	196	27.86	30.04	6.62	.062	2.286	-0.030
1929	3220	176	27.81	29.90	6.60	.077	2.312	-0.037
1930	3280	182	27.75	29.87	6.58	.082	2.321	-0.040
1931	3200	183	27.70	29.86	6.54	.083	2.348	-0.042
1932	3090	175	27.65	29.96	6.57	.075	2.341	-0.038
1933	2870	163	27.59	29.98	6.56	.063	2.340	-0.032
1934	2800	154	27.52	30.08	6.54	.054	2.331	-0.028
1935	2750	158	27.47	30.00	6.54	.057	2.345	-0.030
1936	2700	156	27.42	29.96	6.52	.058	2.337	-0.029
1937	2640	148	27.36	29.83	6.51	.069	2.344	-0.035
1938	2700	152	27.30	29.67	6.52	.084	2.351	-0.043
1939	2650	151	27.25	29.61	6.50	.091	2.355	-0.046
1940	2760	159	27.19	29.36	6.47	.115	2.388	-0.058
1941	2820	168	27.14	29.15	6.42	.143	2.432	-0.072
1942	2950	171	26.94	29.12	6.39	.141	2.446	-0.073
1943	3000	183	26.80	29.15	6.34	.135	2.454	-0.071
1944	3000	173	26.60	29.30	6.37	.114	2.415	-0.059
1945	3000	177	26.42	29.33	6.40	.111	2.394	-0.056
1946	3360	209	26.25	29.01	6.33	.155	2.464	-0.080
1947	3580	225	26.05	28.72	6.29	.184	2.500	-0.095
1948	3420	211	25.87	28.67	6.32	.181	2.498	-0.094
1949	3440	213	25.70	28.60	6.29	.181	2.513	-0.095
1950	3430	213	25.50	28.64	6.36	.109	2.523	-0.100
1951	3480	210	25.33	28.45	6.28	.187	2.522	-0.099
1952	3620	222	25.15	28.35	6.26	.197	2.536	-0.104
1953	3700	225	24.96	28.29	6.25	.202	2.535	-0.106
1954	3810	242	24.78	28.25	6.27	.214	2.548	-0.112
1955	3820	245	24.60	28.22	6.25	.220	2.551	-0.114
1956	3850	254	24.40	28.12	6.23	.232	2.584	-0.123
1957	3930	253	24.23	28.00	6.25	.242	2.587	-0.126
1958	3880	247	24.14	27.93	6.21	.249	2.605	-0.131
1959	3950	257	24.10	27.87	6.19	.263	2.629	-0.140
1960	3910	255	24.06	27.81	6.17	.270	2.636	-0.143
1961	3860	257	24.02	27.77	6.16	.284	2.655	-0.151
1962	3770	255	24.00	27.75	6.11	.291	2.669	-0.155
1963	3690	249	23.97	27.75	6.10	.286	2.672	-0.155
1964	3520	236	23.94	27.80	6.11	.284	2.681	-0.156
1965	3160	212	23.90	27.76	6.14	.276	2.668	-0.151
1966	2820	193	23.86	27.62	6.13	.296	2.700	-0.163
1967	2590	182	23.84	27.41	6.07	.322	2.752	-0.182
1968	2440	170	23.80	27.28	5.98	.341	2.837	-0.209
1969	2410	168	23.76	27.28	5.98	.341	2.837	-0.208

a-Freehand smoothed values of observed modal age of fertility rate.

Source: Data for the calculation of this table are taken from the annual reports of Vital Statistics, 1926-1969, Canada, Dominion Bureau of Statistics.

TABLE A-2.—Parameters Estimated by Elderton's Method: 1926–1969

Year	Modal Age	Y_0	a_1	a_2	m_1	m_2
1926	28.08	177.90	11.10	20.50	0.90	1.67
1927	28.04	175.33	11.12	20.55	0.90	1.66
1928	27.92	174.35	10.99	20.64	0.89	1.66
1929	27.59	171.83	10.59	21.26	0.87	1.75
1930	27.51	175.95	10.43	21.41	0.86	1.77
1931	27.60	173.79	10.68	21.68	0.95	1.93
1932	27.83	166.81	11.03	21.51	0.99	1.93
1933	28.09	161.01	11.61	21.20	1.09	1.99
1934	28.34	151.25	11.93	20.82	1.13	1.97
1935	28.26	149.00	11.99	21.08	1.17	2.05
1936	28.21	146.61	11.88	20,89	1.14	2.01
1937	27.86	143.78	11.30	21.24	1.06	1.99
1938	27.43	147.35	10.64	21.66	0.96	1.95
1939	27.24	144.93	10.34	21.79	0.92	1.93
1940	26.73	153.20	9.69	22.48	0.86	2.00
1941	26.30	159.39	9.14	23.27	0.83	2.12
1942	26.38	167.70	9.34	23.37	0.89	2.22
1943	26.58	172.31	9.62	23.25	0.96	2.33
1944	26.87	169.89	9.97	22.53	0.97	2.19
1945	26.84	168.40	9.88	22.27	0.92	2.07
1946	26.17	194.03	8.98	23.50	0.86	2.25
1947	25.62	209.85	8.34	24.09	0.79	2.29
1948	25.59	199.17	8.46	24.17	0.80	2.30
1949	25.63	201.82	8.65	24.30	0.86	2.41
1950	25.55	199.72	8.55	24.74	0.83	2.41
1951	25.45	205.07	8.54	24.43	0.85	2.43
1952	25.28	214.82	8.34	24.59	0.83	2.45
1953	25.14	219.89	8.15	24.55	0.80	2.41
1954	24.99	226.43	7.96	24.90	0.77	2.41
1955	24.88	228.13	7.76	24.88	0.74	2.38
1956	24.81	232.15	7.76	25.34	0.77	2.53
1957	24.55	236.81	7.51	25.45	0.73	2.47
1958	24.51	236.10	7.49	25.61	0.75	2.55
1959	24.39	241.98	7.35	25.96	0.74	2.62
1960	24.29	241.05	7.19	25.95	0.72	2.61
1961	24.17	239.29	7.01	26.24	0.70	2.64
1962	24.15	236.27	6.91	26.26	0.71	2.68
1963	24.25	231.65	7.07	26.27	0.74	2.75
1964	24.38	220.58	7.27	26.49	0.78	2.85
1965	24.36	196.56	7.38	26.40	0.79	2.81
1966	24.14	176.60	7.20	26.94	0.77	2.90
1967	23.91	165.30	7.01	27.54	0.79	3.09
1968	24.04	159.44	7.39	28.98	0.95	3.71
1969	24.04	157.39	7.39	28.97	0.95	3.71

Acknowledgments

This article is a revised and extended version of the paper appearing in the 1972 *Proceedings of the Social Statistics Section* of the American Statistical Association. We are grateful to the ASA for granting permission to publish this paper.

Particular thanks are expressed to John Kelly as well as to the two anonymous referees for their valuable suggestions regarding both the content and the presentation of the paper. Acknowledgment is also given to Halina Kis and Jeanine Perreault for designing the computer program and to Nancy Pocock and Robert Allen for their technical assistance, all of Statistics Canada.

References

Avery, Roger C. 1970. Graduation of Age-Specific Fertility Rates. A paper presented at the Third Conference on the Mathematics of Population, July 19–24, 1970, University of Chicago.

Canada. Dominion Bureau of Statistics. 1926–1969. Vital Statistics, various years. Ottawa, Canada: Queen's Printer.

Elderton, W. P. 1930. Frequency Curves and Correlation. Cambridge: Cambridge University Press.

Keyfitz, Nathan. 1968. Introduction to the Mathematics of Population. Reading, Mass.: Addison-Wesley Publishing Company.

Mazur, D. Peter. 1963. A Demographic Model for Estimating Age-Order Specific Fertility Rates. Journal of the American Statistical Association 58:774–788.

Mitra, S. 1967. The Pattern of Age-Specific Fertility Rates. Demography 4:894–906.

Murphy, Edmund M., and Dhruva N. Nagnur. 1972. A Gompertz Fit that Fits: Applications to Canadian Fertility Patterns. Demography 9:35–50.

Ryder, Norman B. 1969. The Emergence of a Modern Fertility Pattern; United States, 1917–1966. Pp. 99–123 in S. J. Behrman, Leslie Corsa, Jr., and Ronald Freedman (eds.), Fertility and Family Planning. Ann Arbor: University of Michigan Press.

Siegel, Jacob S., and Donald S. Akers. 1969. Some Aspects of the Use of Birth Expectations Data from Sample Surveys for Population Projections. Demography 6:101–115.

Stone, Leroy O. 1970. Parametric Approaches of the Age Distribution of Cohort Total Fertility for Projections. Unpublished paper. Ottawa, Canada: Dominion Bureau of Statistics.

Tekse, K. 1967. On Demographic Models of Age-Specific Fertility Rates. Statistisk Tidskrift (Stockholm), Series III. Statistical Review, III 5:189–207.

Wicksell, S. D. 1931. Nuptiality, Fertility and Reproductivity. Pp. 125–157 in Skandinavisk Aktuarietidskrift.

CENSUS FIELD PAPERS AND MEMORANDA ISSUED

Reprinted Publications (since 1965)

Census Technical Papers (1965-1968)

Technical Memoranda (1965-1968)

Analytical and Technical Memoranda (since 1969)

Lists of papers published in above series are available from:

Technical Co-ordinator (Research)
Census Field
Statistics Canada
Ottawa, Ontario
K1A OT7